*Journal of Applied Economics, Vol. I, No. 2 (Nov. 1998), 413-453*

# DO COMMUNITY CHARACTERISTICS INFLUENCE ENVIRONMENTAL OUTCOMES? EVIDENCE FROM THE TOXICS RELEASE INVENTORY

**SEEMA ARORA***
*Vanderbilt University*

and

**TIMOTHY N. CASON***
*Purdue University*

This research uses neighborhood characteristics (at the zipcode level) in 1990 to explain toxic releases in 1993. It combines the Toxics Release Inventory data with demographic data from the 1990 US Census. We first analyze the location of manufacturing facilities in a particular neighborhood using a sample selection model, and then estimate the relationship between releases in 1993 and the demographic characteristics of the neighborhood in 1990. We conduct the analysis for the entire US as well as for different geographic regions to study regional differences in determinants of environmental outcomes. Releases in non-urban areas of the southeastern US exhibit a pattern suggesting that race might be an important determinant of release patterns. Economic characteristics of neighborhoods (such as income levels and unemployment) also affect releases. Our variables that proxy the propensity for communities to engage in political action exert greater influence on environmental outcomes in non-urban areas.

---

* S. Arora is a professor at the Owen Graduate School of Management and T. N. Cason is a professor at the Department of Economics and Krannert School of Management. We have benefited from helpful comments provided by seminar participants at UC-Santa Barbara, the University of Southern California and Resources for the Future and conference participants at the European Agricultural and Resource Economists Meeting in Lisbon, Portugal. We would like to thank David Austin, Dallas Burtraw, Mark Cohen, Brian Kropp, Eduardo Ley, Vai-Lam Mui, Wallace Oates, Ian Parry, Hilary Sigman, Jeff Wagner, Margaret Walls, and Chris Wernstedt. We retain responsibility for any errors. Arora gratefully acknowledges financial support from the Owen Graduate School of Management Dean's Fund for Summer Research.

## I. Introduction

The traditional methods of command and control regulation have been ineffective at worst and costly at best. Recognizing the need to make regulations more flexible, in the past decade Congress and regulators have started to favor innovative and more market based approaches to regulation. The use or proposed use of tradable permits for controlling acid rain and more recently for mitigating global warming exemplify this trend toward more flexible and market oriented approaches. The use of public information is yet another innovative environmental policy tool. While economists pushed for the adoption of tradable permits by appealing to its cost effectiveness, policy makers adopted public information disclosure without prodding by economists. Congress was inspired by an industrial accident in Bhopal, India when it passed the Emergency Planning and Community Right-to-Know Act (EPCRA) in 1986. EPCRA requires all manufacturing facilities to make public their releases of over 320 toxic chemicals. The underlying premise of public disclosure as an environmental policy tool is that public knowledge of pollution can engender effective and informed participation by various constituencies to exert pressure on facilities to improve their environmental performance.

Public knowledge of environmental data can be used by consumers to boycott products, or it may be used by investors to penalize large polluters (Hamilton, 1995b; Konar and Cohen, 1997). Neighborhood characteristics may also influence enforcement actions by regulators.[1] This paper analyzes the role of communities in influencing environmental outcomes. We examine the potential impact of public disclosure on the environmental performance of facilities by studying how community characteristics such as race and gender, economic status and variables expected to capture political action influence subsequent toxic releases. A number of studies have concentrated

[1] For an informational model of Occupational Saftey and Health Administration enforcement, see Scholz and Gray (1997).

on the relationship between race and environmental outcomes to determine the extent of environmental injustice.[2] In the present paper we find evidence of environmental injustice, and we also examine the effects of other community characteristics in influencing environmental results.

We combine the Toxics Release Inventory data with demographic data from the 1990 US Census. We use neighborhood characteristics (at the zipcode level) to explain toxic releases in 1993, controlling for releases in 1990. Releases in a particular year are determined simultaneously with the demographic characteristics of a neighborhood, and they change over time for a variety of reasons—including facility relocation, expansion and downsizing, as well as in response to community characteristics. Because the releases in 1993 are determined after the demographic characteristics were determined in 1990, it is reasonable to treat the demographic characteristics as exogenous with respect to these later releases.

We first analyze the location of manufacturing facilities in a particular neighborhood using a sample selection model. This first stage relates the likelihood that a neighborhood experiences any toxic releases to the characteristics of that neighborhood. We then attribute the level of emissions in 1993 to the demographic and socio-economic characteristics of the neighborhood in 1990. We conduct the analysis for the entire United States as well as specific geographical regions.

The analysis captures three distinct aspects of the communities to assess the role that each plays in influencing environmental outcomes. First we consider the racial, immigrant and gender composition of neighborhoods. Our results indicate that a larger percentage of non-white residents may be

[2] For previous research, see Anderton et al. (1994), Bryant and Mohai (1992), Bullard (1983 and 1990), Goldman and Fritton (1994) and Been (1994).

associated with a higher level of releases in the southeastern states, primarily in non-urban zipcodes.[3] We also examine the relationship between economic characteristics and environmental outcomes. Economic factors (such as median income and unemployment rates) have a significant impact on toxic releases, particularly in the southeastern states. Finally, we examine variables expected to be associated with the political activity and preferences of the community and its ability to collectively oppose firms that may harm the local environment. While we use voter turnout data and data on environmental initiative voting for California, for the rest of the US we use demographic variables as proxies to represent a community's propensity for collective action and its political preferences. These variables appear to influence environmental outcomes mainly in non-urban areas.

## II. Theoretical Framework and Hypotheses Construction

Hamilton (1995a) presents a careful description of three alternative explanations for pollution patterns resulting from capacity expansion plans for commercial hazardous waste facilities, and we adopt his framework to motivate our empirical hypotheses. The three explanations are (1) race/gender related, (2) the Coase theorem and (3) the theory of collective action (Olson, 1965). In the first explanation, facility owners and operators consider the race and gender composition of neighborhoods and increase releases in neighborhoods with a greater minority (and perhaps immigrant) population, or with a greater fraction of female-headed households. In its pure form, this leads to greater releases in some neighborhoods that otherwise (from a pure profit-maximizing standpoint) would not experience greater releases.

[3] As documented in Section IV. D, only in the southeastern states do racial minorities commonly represent a large proportion of total residents in non-urban areas.

Alternatively, in a world without transaction costs the Coase theorem suggests that releases will increase in neighborhoods in which the releases will do the least damage. According to this hypothesis, releases will be greater in neighborhoods with lower rent. Higher incomes may also increase the costs of increased releases in a given neighborhood.[4] Rental values and income levels are correlated with education and race, so releases could increase in minority neighborhoods merely because they affect lower-valued property and lower wage earners. Our analysis controls for rental values and income in an attempt to sort out these alternative explanations.[5]

Finally, firms may decide to increase releases in a given neighborhood because they face less (political) collective action in that neighborhood. Residents in different neighborhoods vary in their ability to overcome free-rider problems and engage in collective action. Again, this could result in outcomes that appear similar to the race/gender-related explanation if, for example, minority or immigrant neighborhoods are less politically active. To distinguish between these explanations we include some variables that are likely to affect incentives to engage in collective action (such as the fraction of households with children); and in a model based on California data only we include some direct measures of political action and environmental preferences—voter turnout and vote results on an environmental initiative. While we can use voting data for California, due to data limitations for other regions (discussed below) we rely on a combination of demographic variables to proxy for collective action.

---

[4] From the polluter's perspective, higher property values and incomes increase the damage from releases because in litigation, injured parties could recover damages based on reduced property values. In the case of adverse health impacts that limit work ability the injured parties could recover lost income.

[6] This is a different point than stated by Been (1994). She argues that releases in a neighborhood decrease property values, which then attract minority populations. Econometrically, this suggests that neighborhood characteristics may be endogenous to the determination of releases. This is precisely why we use 1990 characteristics to explain 1993 releases; see Section 3.4.

Strong correlations exist between many of our explanatory variables, which creates a classic multicollinearity problem. This problem has the potential to cause incorrect statistical inferences regarding individual coefficient estimates. This potential arises because although individual coefficient estimates are unbiased, variance estimates are inflated due to the multicollinearity. To sidestep this problem we focus on joint tests of significance to test the three alternative hypotheses. In particular, we employ the Wald test in a series of hypothesis tests of the form $H_0$: $Rb=r$, where $R$ is a matrix that creates a joint test that specific elements in the parameter vector $b$ are all equal to zero ($r$ is a vector of zeros). We choose three different $R$ matrices to test each of the three explanations described above.

To summarize, these alternative theories predict that only certain variables should explain toxic releases. The race/gender hypothesis posits the null th at factors such as race, gender and the foreign-born composition of a neighborhood do not predict releases. Rejection of the null implies that these factors are important and supports the race/gender hypothesis. The economic (Coase theorem) hypothesis postulates the null that economic factors such as income levels, rental values, vacancy rates, unemployment rates and the proportion of poor households do not explain changing release patterns. Rejection of this null supports what we shall refer to as the economic/Coasian explanation for changing release patterns. Lastly, the political/collective action hypothesis posits the null that variables related to the political action propensity of local residents do not predict releases. In addition to voter turnout and expressed preferences through environmental initiative voting (for California only), we include variables such as age, education and the number of households with children.[6] These factors can be reasonably expected to influence the

[6] Recall the incident at Love Canal, where an elementary school was built on a toxic dump. That caused a public outcry when the chemicals started seeping from the walls and affecting children.

incentives and tendency to engage in political action (e.g., see Filer et al., 1993).[7] Rejection of this political/collective action null supports the hypothesis that such variables associated with the political activity of local residents influence environmental outcomes.

We focus on hypothesis tests for these three sets of variables as a group, and then also interpret the significant individual variable effects. We recognize that our classification of variables under the different hypotheses is not exact. For example, the proportion of foreign-born residents may be associated primarily with the race/gender hypothesis, but it may also be considered a factor that influences the extent of community activism. Our presentation of individual coefficient estimates permits the reader to assess the implications of alternative groupings.[8]

## III. Data and Model Specification

We combine the Toxics Release Inventory with the US Bureau of the Census data and determine the relationship between the releases in a particular zipcode and demographic attributes of that zipcode. We use data for nearly 30,000 zipcodes, including all zipcodes with residential population according to the US Census.

---

[7] Filer et al. (1993) use variables such as education, age and income to explain voter turnout. In the set of political/collective action variables, we also include several factors that potentially affect or reflect local environmental preferences. We include the percentage of residents who carpool because carpooling for some may represent a contribution to a community public good or pro-environmental preferences. The percentage of residents employed in manufacturing industries and the percentage of residents who rent rather than own their residences are also included in the set of political action variables because these variables could influence the incentives for residents to oppose expansions in local manufacturing facilities.

[8] We should also note that because of the inexact variable classification and the multicollinearity present in these demographic data, our Wald tests of joint significance of each set of variables could be sensitive to alternative groupings.

### A. The Toxics Release Inventory

Title III of the Superfund Amendments and Reauthorization Act (1986) requires manufacturing establishments (Standard Industrial Classification (SIC) 20-39) to report their releases and transfers of 320 toxic chemicals. The Act requires facilities that manufacture or process more than 25,000 pounds or use more than 10,000 pounds of any of the reportable chemicals to submit a TRI report [EPA (1992)]. Our main results aggregate air, land, water and underground injection releases, and do not include toxic chemical transfers. [Section IV. E. briefly discusses models estimated for toxic transfers and releases disaggregated by release medium.] Arora and Cason (1995) compare two methods of chemical aggregation—one weighting all chemicals equally and another that accounts for the chemicals' different toxicity. Most of the toxic chemicals that are widely used have similar toxicity [EPA(1989)], so the results were not sensitive to the weighting scheme.[9] Therefore, here we simply aggregate the chemicals and employ equal weights.[10]

In addition to the environmental data, each facility reports its location, primary SIC code and parent company. We employ the zipcode of the facility location to merge these data with the Census data. Note that our measure of environmental outcomes is based on releases and not exposures. Exposures differ from releases due to the geographic dispersion of households and releases within each zipcode. We do not attempt to analyze exposures here as it would entail very elaborate mappings using the census tract and a geographical information system. Given the scope of our study (for the entire US) this exercise is prohibitively expensive. Note also that since the analysis

---

[9] Indeed, EPA has not assigned risk scores to many of the less toxic chemicals on the TRI list, which makes differential weighting problematic.

[10] A limitation of the TRI data set is that it is self reported so there may be an incentive to under-report the releases. There may also exist an incentive to over-report if firms expect to be rewarded for improvements relative to a baseline emission level. Nevertheless, at present it is the best available dataset that provides a comprehensive analysis of toxic release patterns for the entire US. While there is some non-reporting it appears due to ignorance and not evasion (Brehm and Hamilton, 1996).

is conducted at the zipcode rather than at the firm level, it is not possible to control for industry since multiple facilities (from multiple industries) exist in many zipcodes.

### B. The Census Data

The Sourcebook of Zip Code Demographics compiles the 1990 US Census separately for every residential zipcode. Table 1 summarizes the variables we employ. All variables are for 1990 unless noted otherwise. Using the zipcode level of aggregation is most straightforward and practical given this broad-based study of the entire US. Some spatial correlation of releases and demographic characteristics undoubtedly exists, but numerically adjacent zipcodes are often not adjacent geographically. Therefore, accounting for this correlation would also require a detailed geographic information system. This is more practical for less broad studies, such as the analysis of health risks in Pennsylvania's Allegheny County conducted by Glickman and Hersh (1995).

### C. Additional California Variables

We present results in Section IV. C. based on California zipcodes, after adding two variables that that we obtained only for California—voter turnout and vote outcomes on a specific ballot proposition. These variables are intended to capture the political activity and environmental preferences of residents of different areas of the state. Unlike the other zipcode-specific demographic and economic characteristics described above, these data are provided at the county level.[11]

[11] It would be possible, in principle, to collect voter turnout data for every state; unfortunately, such data are compiled at the state rather than federal level. Moreover, we have not identified a compilation of national voter turnout data with zipcode or numerical county identifiers that are suitable for merging with the zipcode or county identifiers on the census database. The California Secretary of State also compiles voting data at different levels of aggregation—such as by Congressional district—but they are not compiled by zipcode. For our analysis, we merge the county-based voting data with the zipcode-level demographic and socio-economic data. We thank John Matsusaka for generously providing these voting data.

**Table 1. Description of the CENSUS data**

The Sourcebook of Zip Code Demographics provides data on all residential neighborhoods in the region. All variables are for 1990, in 1990 $.

| Variable | Definition |
|---|---|
| FEMHEAD | Percentage of family households with a female as the head of the household |
| PCTFORN | Percentage of foreign born residents |
| PCTNONWT | Percentage of non-white residents (Black, American Indian, Asian/Pacific Islander, Other) |
| PCTASIAN | Percentage of residents classified as Asian/Pacific Islander |
| PCTNONWA | Percentage of non-white and non-Asian residents (Black, American Indian, Other) |
| VACANT | Percentage of housing units that are vacant. Includes housing units that were temporarily occupied at the time of the census; seasonal or recreational units, units for sale or rent, units rented or sold but not occupied, and new units not occupied. |
| MDINCOME | Median household income. The median has been computed from the nine intervals in the reported distribution of income. |
| POOR | Percentage of residents living in poverty. Poverty status is calculated in 1989. Poverty thresholds are calculated from the number of persons in the family and the number of related children under 18 years. The average threshold for a family of four in 1989 was $12,674; for two persons it was $8,076. |
| MEDROHU | Median rent paid in renter occupied housing units (dollars per month) |
| UNEMP | Unemployment rate (in percent) |
| BACH | Percent of population (over 25 years of age) with bachelor's degree |
| CARPOOL | Percentage of workers sixteen years and older who journey to work by carpool |
| HHWKIDS | Percentage of family households with children (below 18 years of age) |
| MANU | Percentage of workers employed in manufacturing industries |
| MEDAGE | Median age of residents |
| RENTPCT | The percent of occupied housing units that are renter occupied. Contract rent is the monthly amount, regardless of any utilities, furnishings, or fees, that may be included. These renter-occupied units exclude single family homes on more than 10 acres and renter units that are occupied without payment of cash rent. |
| TOTPOP | The total number of residents in an area, where residence refers to the "usual place" where a person lives, which is not necessarily the legal residence. |
| PCTURB | Percentage of residents living in an urban area. Urban includes population of places with at least 2500 persons and urbanized area. Urbanized area consists of one or more places with a minimum population of 50,000 people plus adjacent area with a density of 1000 persons per square mile. |

We employ voter turnout from 1990, the same year as the census data. The turnout measure is the total votes cast in the county in the 1990 general election, as a percentage of the total 1990 population in the county. Traditional measures of voter turnout use either eligible or registered voters in the denominator. We chose total population for our denominator so that our measure captures not only the political activity of the residents, but also level of enfranchisement of the population. Our version differs from traditional measures because the proportion of children, immigrants, and others ineligible to vote varies across counties. Our logic is that the political influence of a population declines if either (a) the eligible voters in that population tend to vote less often or (b) more members of that population are ineligible to vote. The measure we construct combines these two components of political activity.

The proposition we chose to represent environmental preferences is Proposition 128, popularly known as "Big Green," which was defeated in the 1990 general election. The most notable feature of the proposition was a ban on the use of pesticides that cause cancer or reproductive harm, which would have eliminated about 350 chemicals (out of about 2,300 currently in use). The initiative was also wide-ranging, including a ban on new offshore oil drilling, increased water quality standards, $300 million in bonds to buy redwoods, and a proposal to reduce greenhouse gas emissions by 40 percent. Clearly, an increase in the proportion of voters voting for proposition 128 in a region indicates more pro-environment preferences in that region.[12]

### D. Model Specification

Our goal is to explain the toxic chemical releases in 1993 using the socio-economic characteristics and 1990 releases of zipcode neighborhoods. Most prior research investigating the relationship between demographic variables and environmental outcomes fails to recognize that the neighborhood

---

[13] See Kahn and Matsusaka (1997) for a comprehensive analysis of voting behavior on a large sample of California initiatives.

characteristics and environmental outcomes are determined simultaneously. A facility locates in an area, increasing the environmental risk and causing the land and housing values of that area to decline. Residents that choose to live in that area may either place a low value on the environment or may have a low income that limits their ability to locate in a less environmentally degraded area. Our strategy to avoid this endogeneity problem is to use 1990 demographic characteristics to explain releases after 1990. Increases in releases occur from new facilities or expansion of existing facilities after 1990, so the 1990 demographic characteristics are most likely exogenous to these post-1990 firm decisions. We do acknowledge, however, that our results are still subject to some (we believe minor) endogeneity bias if residents are located in a given neighborhood in 1990 based on expectations of how releases will change after 1990.[13]

An immediate problem that arises in constructing the dependent measure of toxic releases is that many neighborhoods do not have any toxic chemical releases in either 1990 or 1993. In particular, 72 percent of the nearly 30,000 zipcodes with demographic data experienced no toxic chemical releases according to the TRI in these years. Simply excluding these zipcodes from our analysis would lead to a potentially significant sample selection bias, since these zero-release neighborhoods are obviously not a random sample of neighborhoods. We therefore employ a two-stage maximum likelihood sample selection model so that our estimates of the releases equation account for the non-random selection of the neighborhoods with any toxic chemical releases (Heckman, 1979). The first stage estimates a probit model, with the dependent variable equal to 1 if the neighborhood experienced any toxic releases in 1990 or 1993 (and 0 otherwise). The second stage estimates our main model (with 1993 releases as the dependent variable), adding the

[13] Another approach might be to determine environmental performance by measuring something like the level of releases per $1000 in value-added for these manufacturing facilities. This would involve merging detailed data from the manufacturing census, an ambitious avenue of inquiry that we leave for future research.

estimated likelihood of any releases for that zipcode calculated from the first stage (or what is commonly referred to as the inverse Mill's ratio).

The second econometric issue that arises is heteroscedasticity. Zipcode boundaries are designed to facilitate the delivery of mail rather than group the population into roughly equal-sized neighborhoods; consequently, the number of residents in each zipcode varies considerably.[14] More populous zipcode neighborhoods were more likely to experience toxic releases, and a Breusch and Pagan (1980) Lagrange multiplier test strongly rejects homoscedasticity at better than the p=0.001 significance level.[15] To account for this heteroscedasticity in the estimates we assume that the standard deviation in each observation is proportional to the residential population of the zipcode neighborhood. This assumption is translated into the econometric estimation by weighting each observation by the inverse of the square root of residential population.

Table 2 presents summary statistics for the analysis variables. Column (1) presents a summary of the socio-economic characteristics of the zipcode neighborhoods with no toxic releases in either 1990 or 1993. Column (2) presents this same information for the neighborhoods with positive releases in either 1990 or 1993.

## IV. Results

Table 3 presents total toxic releases reported in the TRI for 1990 and 1993. Nationally, releases declined by 6.5 percent. The table also shows that the decline in releases was more modest in the southeastern U.S., a

---

[14] A number of entirely industrial or commercial zipcodes have no residents, so they have no demographic data and cannot contribute to our analysis. The most populous zipcode had 112,046 residents.

[15] This test statistic is simply one-half of the explained sum of squares in the regression of $u_i = e_i^2/(e'e/N) - 1$ on the vector of explanatory variables. We conducted this test based on the second stage regression that includes the inverse Mill's ratio to account for sample selection.

**Table 2. Summary Statistics**

| | | zipcodes with no releases in 1990 and 1993 | zipcodes with positive releases in 1990 or 1993 |
|---|---|---|---|
| | | (N=21215) | (N=8122) |
| FEMHEAD | Median | 9.90 | 13.70 |
| (Percentage of Female- | Mean | 11.29 | 16.06 |
| Headed Family Households) | Std. Dev. | 6.44 | 8.85 |
| PCTFORN | Median | 1.00 | 2.00 |
| (Percentage of Foreign- | Mean | 2.94 | 4.92 |
| Born Residents) | Std. Dev. | 5.79 | 7.90 |
| PCTNONWT | Median | 2.50 | 6.70 |
| (Percentage of Non-White | Mean | 10.56 | 15.90 |
| Residents) | Std. Dev. | 18.03 | 20.61 |
| VACANT | Median | 11.80 | 7.40 |
| (Percentage of Housing | Mean | 16.32 | 9.35 |
| Units that are Vacant) | Std. Dev. | 13.76 | 7.40 |
| MDINCOME | Median | 24.06 | 26.84 |
| (Median Household Income, | Mean | 26.05 | 28.69 |
| in Thousands) | Std. Dev. | 9.66 | 9.63 |
| POOR | Median | 13.00 | 11.50 |
| (Percentage of Residents | Mean | 14.92 | 13.60 |
| living in Poverty) | Std. Dev. | 9.70 | 9.42 |
| MEDROHU | Median | 216.00 | 287.00 |
| (Median Rent in Renter-Occ. | Mean | 265.72 | 327.64 |
| Housing Units, per Month) | Std. Dev. | 158.28 | 151.10 |
| UNEMP | Median | 5.70 | 5.80 |
| (Unemployment Rate in | Mean | 6.74 | 6.59 |
| Percent) | Std. Dev. | 4.40 | 3.54 |
| BACH | Median | 11.10 | 13.20 |
| (Percentage of Population | Mean | 14.51 | 16.41 |
| over 25 with Bachelors Degree) | Std. Dev. | 10.92 | 10.58 |
| CARPOOL | Median | 14.90 | 13.80 |
| (Percentage of Labor Force over | Mean | 15.74 | 14.52 |
| 16 years old who carpool) | Std. Dev. | 5.92 | 4.72 |
| HHWKIDS | Median | 50.00 | 51.20 |
| (Percentage of Households | Mean | 49.94 | 51.43 |
| with Children < 18 years) | Std. Dev. | 8.50 | 7.02 |

**Table 2. Summary Statistics (continued)**

| | | zipcodes with no releases in 1990 and 1993 | zipcodes with positive releases in 1990 or 1993 |
|---|---|---|---|
| MANU | Median | 15.30 | 20.30 |
| (Percentage of Labor Force | Mean | 16.84 | 21.53 |
| in Manufacturing Industries) | Std. Dev. | 10.38 | 9.46 |
| MEDAGE | Median | 34.70 | 33.60 |
| (Median Age of Residents) | Mean | 34.90 | 33.52 |
| | Std. Dev. | 4.94 | 3.76 |
| RENTPCT | Median | 21.30 | 28.20 |
| (Percentage of Residents | Mean | 24.62 | 31.72 |
| Renting Primary Residence) | Std. Dev. | 13.80 | 15.05 |
| TOTPOP | Median | 1.63 | 13.37 |
| (Residential Population | Mean | 5.04 | 17.43 |
| in Thousands) | Std. Dev. | 9.15 | 14.97 |
| SUMREL90 | Median | 0.00 | 36.74 |
| (Toxic Releases reported | Mean | 0.00 | 408.82 |
| for 1990, in Thousands) | Std. Dev. | 0.00 | 2966.61 |
| SUMREL93 | Median | 0.00 | 26.30 |
| (Toxic Releases reported | Mean | 0.00 | 382.45 |
| for 1993, in Thousands) | Std. Dev. | 0.00 | 3110.68 |
| PCTURB | Median | 0.00 | 74.30 |
| (Percentage of Residents | Mean | 21.69 | 62.28 |
| living in Urban Areas) | Std. Dev. | 36.78 | 38.10 |

**Table 3. Total Toxic Releases Reported in the Toxics Release Inventory (in Millions of Pounds)**

| Year | Entire U.S. | South | Non-South |
|---|---|---|---|
| 1990 | 3,905 | 1,518 | 2,387 |
| 1993 | 3,653 | 1,491 | 2,161 |
| Percentage Change | -6.5% | -1.8% | -9.5% |

region comprised of 11 states (Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee and Virginia). This difference, in part, motivated us to estimate models separately for this region. Section IV.B. presents these regional estimates, following the full sample estimates in Section IV.A.

### A. Full Sample Estimates

Panel A of Table 4 contains the probit sample selection parameter estimates, and Panel B of Table 4 contains the parameter estimates that explain the toxic releases in 1993. Column (1) presents estimates based on all zipcodes in the United States with any residential population. No causality should be inferred from the Panel A sample selection estimates; as discussed above, the existence of toxic releases in a particular neighborhood undoubtedly influences the decision of many residents to locate in that neighborhood, and therefore partially explains its socio-economic characteristics. The sample selection equation is used merely to retrieve the inverse Mill's ratio, so in this discussion we focus on the Panel B estimates.[16]

Due to differences in state regulations as well as other differences due to economic conditions, releases could differ across states. We therefore include 49 state dummy variables, but suppress them in the tables to conserve space. The omitted dummy variable is for the most populous state (California). Fourty-six of the 49 state dummy variables are not significantly different from zero (at the 5-percent level), indicating that fixed state effects are usually not important.[17] The remaining estimates in Table 4 Panel B are marginal

[16] As shown at the bottom of Table 4, Panel B, the inverse Mill's ratio sample selection term is never significant. This suggests that any sample selection bias is probably small—which we confirm with ordinary least squares estimates shown in the Appendix. We nevertheless focus on the sample selection model shown in Table 4 because it is reasonable to expect a selection bias, at least in theory.

[17] The three significant state dummy variables are for Kansas (estimate=-503.8), Louisiana (estimate=1234.3) and Utah (-1194.3).

**Table 4. Estimation Results, by Region (Full Sample)**

**Panel A. Stage 1. Sample Selection Probit Equation (Dependent Variable is = 1 if Any Releases Reported)**

| | U.S. | | South | | Non-South | | California | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Estimate | Std. Error | Estimate | Std. Error | Estimate | Std. Error | Estimate | Std. Error |
| Constant | -0.882*** | (0.194) | -0.364 | (0.410) | -1.216*** | (0.241) | 0.747 | (0.725) |
| FEMHEAD | 0.031*** | (0.003) | 0.051*** | (0.006) | 0.025*** | (0.003) | 0.007 | (0.012) |
| PCTFORN | -0.014*** | (0.002) | -0.026*** | (0.004) | -0.009*** | (0.002) | 0.004 | (0.006) |
| PCTNONWT | -0.007*** | (0.001) | -0.004** | (0.002) | -0.010*** | (0.001) | — | — |
| PCTASIAN | — | — | — | — | — | — | -0.005 | (0.006) |
| PCTNONWA | — | — | — | — | — | — | -0.004 | (0.004) |
| VACANT | -0.004*** | (0.001) | -0.003 | (0.003) | -0.004*** | (0.001) | -0.017** | (0.007) |
| MDINCOME | 0.007*** | (0.002) | 0.010* | (0.005) | 0.004* | (0.003) | -0.007 | (0.008) |
| POOR | -0.004* | (0.002) | -0.016*** | (0.004) | 0.002 | (0.002) | 0.003 | (0.012) |
| MEDROHU | -0.001*** | (0.000) | -0.001* | (0.000) | -0.001*** | (0.000) | 0.000 | (0.001) |
| UNEMP | -0.027*** | (0.004) | -0.023*** | (0.009) | -0.027*** | (0.005) | 0.031 | (0.020) |
| BACH | -0.014*** | (0.001) | -0.023*** | (0.003) | -0.011*** | (0.002) | -0.007 | (0.006) |
| CARPOOL | -0.019*** | (0.002) | -0.035*** | (0.005) | -0.013*** | (0.003) | -0.004 | (0.012) |
| HHWKIDS | -0.002 | (0.002) | -0.006 | (0.004) | -0.001 | (0.002) | -0.012** | (0.006) |
| MANU | 0.039*** | (0.001) | 0.029*** | (0.002) | 0.046*** | (0.001) | 0.042*** | (0.006) |
| MEDAGE | -0.019*** | (0.003) | -0.016** | (0.007) | -0.017*** | (0.004) | -0.051*** | (0.014) |
| RENTPCT | 0.002 | (0.001) | 0.003 | (0.002) | 0.002* | (0.001) | -0.007* | (0.004) |
| TURN90 | — | — | — | — | — | — | 0.001 | (0.008) |
| PCT4_128 | — | — | — | — | — | — | -0.008 | (0.007) |
| TOTPOP | 0.028*** | (0.001) | 0.049*** | (0.003) | 0.024*** | (0.001) | 0.019*** | (0.003) |
| PCTURB | 0.012*** | (0.000) | 0.005*** | (0.001) | 0.013*** | (0.000) | 0.009*** | (0.002) |
| Log-Likelihood | | -12287 | | -3045 | | -9104 | | -658 |

**Table 4 (continued). Estimation Results, by Region (Full Sample)**

**Panel B: Stage 2: Dependent Variable is 1993 Releases (in Thousands of Pounds)**

| | U.S. | | South | | Non-South | | California | |
|---|---|---|---|---|---|---|---|---|
| | (1) Estimate | (2) Std. Error | (3) Estimate | (4) Std. Error | (5) Estimate | (6) Std. Error | (7) Estimate | (8) Std. Error |
| Constant | -2095.80*** | (688.63) | -5353.70*** | (1679.70) | 74.53 | (538.30) | 187.27 | (1655.40) |
| | *Race/Gender Variables* | | | | | | | |
| FEMHEAD | 14.36 | (14.80) | -9.64 | (49.51) | -1.65 | (10.62) | -49.19 | (57.06) |
| FEMHEDSQ | -0.65*** | (0.21) | -1.12* | (0.61) | -0.03 | (0.16) | 1.02 | (1.08) |
| PCTFORN | -8.82 | (6.07) | 11.89 | (22.47) | -7.14* | (4.13) | -8.48 | (13.71) |
| PCTNONWT | -11.73** | (5.10) | -38.00*** | (11.73) | -0.56 | (4.30) | — | — |
| PCTNWTSQ | 0.27*** | (0.06) | 0.79*** | (0.14) | -0.02 | (0.05) | — | — |
| PCTASIAN | — | — | — | — | — | — | -12.52 | (29.68) |
| PCTASQ | — | — | — | — | — | — | 0.18 | (0.68) |
| PCTNONWA | — | — | — | — | — | — | -9.42 | (18.80) |
| PCTNWASQ | — | — | — | — | — | — | -0.05 | (0.21) |
| | *Economic Variables* | | | | | | | |
| VACANT | -5.85* | (3.13) | -20.67* | (12.22) | -2.67 | (2.11) | -32.91 | (21.59) |
| MDINCOME | 109.59*** | (34.32) | 370.62*** | (137.54) | -11.75 | (24.44) | 81.46 | (136.76) |
| MEDINCSQ | -2.41*** | (0.91) | -8.66** | (4.35) | 0.29 | (0.63) | -2.53 | (3.81) |
| MEDINCCU | 0.02** | (0.01) | 0.07 | (0.04) | 0.00 | (0.01) | 0.02 | (0.03) |
| POOR | -7.60 | (11.86) | -43.70 | (43.06) | 5.36 | (8.18) | 35.31 | (47.57) |
| POORSQ | 0.77*** | (0.23) | 2.37*** | (0.77) | -0.07 | (0.18) | -0.71 | (0.95) |
| MEDROHU | -0.64* | (0.38) | -0.86 | (1.35) | -0.22 | (0.26) | 0.11 | (1.06) |
| UNEMP | 33.78 | (21.62) | 241.37*** | (75.87) | 3.13 | (15.78) | -61.50 | (99.23) |
| UNEMPSQ | -3.05*** | (0.91) | -16.01*** | (3.69) | -0.23 | (0.66) | 3.18 | (3.43) |

**Table 4 (continued). Panel B**

| | U.S. | | South | | Non-South | | California | |
|---|---|---|---|---|---|---|---|---|
| | (1) Estimate | (2) Std. Error | (3) Estimate | (4) Std. Error | (5) Estimate | (6) Std. Error | (7) Estimate | (8) Std. Error |
| | *Political/Collective Action Variables* | | | | | | | |
| BACH | -1.28 | (4.43) | 14.13 | (17.31) | -5.51* | (2.82) | -27.67* | (14.18) |
| CARPOOL | 24.01 | (20.46) | 109.13 | (73.34) | 0.07 | (14.11) | -41.24 | (51.04) |
| CARPOLSQ | -0.50 | (0.54) | -2.22 | (1.69) | 0.28 | (0.40) | 0.64 | (0.93) |
| HHWKIDS | 7.59 | (20.33) | 34.92 | (56.74) | -2.96 | (15.29) | 1.82 | (53.85) |
| HHWKIDSQ | -0.12 | (0.21) | -0.49 | (0.60) | 0.00 | (0.16) | 0.03 | (0.56) |
| MANU | -2.29 | (8.88) | -27.11 | (17.08) | 2.31 | (6.98) | 40.58 | (31.35) |
| MEDAGE | 6.57 | (8.96) | 9.24 | (27.21) | 2.45 | (6.21) | -55.91 | (54.80) |
| RENTPCT | 3.64 | (2.93) | 17.45* | (10.23) | 1.97 | (2.03) | 3.24 | (11.10) |
| TURN90 | — | — | — | — | — | — | -18.99 | (15.86) |
| PCT4_128 | — | — | — | — | — | — | 2.91 | (14.89) |
| | *Control Variables* | | | | | | | |
| TOTPOP | 4.82 | (5.16) | -13.17 | (21.15) | 5.44* | (3.08) | 16.57 | (13.85) |
| 1990 Releases | 0.94*** | (0.01) | 1.07*** | (0.01) | 0.61*** | (0.01) | 0.34*** | (0.08) |
| PCTURB | -4.50 | (3.84) | -12.08 | (7.84) | 1.31 | (2.84) | -9.11 | (17.05) |
| PCTURBSQ | 0.05** | (0.02) | 0.01 | (0.07) | 0.01 | (0.02) | 0.17 | (0.11) |
| Inverse Mill's Ratio | 217.70 | (310.98) | -541.64 | (770.36) | 128.26 | (212.07) | 1535.20 | (1199.80) |
| Number of Observations | | 29332 | | 6691 | | 22641 | | 1501 |
| Adjusted $R^2$ | | 0.735 | | 0.813 | | 0.529 | | 0.044 |

Notes: * denotes significantly different from zero at the 10 percent level; ** denotes significantly different from zero at the 5 percent level; *** denotes significantly different from zero at the 1 percent level (all two-tailed tests).

within-state impacts of the demographic characteristics because across-state differences are captured by the fixed effect state dummies.

We have no prior that suggests only a linear relationship between any of our explanatory variables and releases, and some case studies (Bullard, 1983; GAO, 1983) have found negative environmental outcomes only when certain factors (such as the non-white population) are very high in the local population. For these reasons we include squared terms for many of the variables. Preliminary estimates indicated that no significant non-linear relationships for certain variables, so Table 4 presents estimates without squared terms for those variables when the preliminary estimates indicated a squared term coefficient that was only a small fraction of its standard error. We also included cubic terms in preliminary regressions; these were all insignificant except for median income, which we therefore include (MEDINCCU). The cubic functional form for median income permits a sufficiently non-linear relationship to represent the inverse U-shaped environmental Kuznets' curve identified by Grossman and Krueger (1995) based on a panel of cities in different countries. The interpretation of the curve inverse U-shaped environmental Kuznets' curve (relating income and an environmental indicator) is that as economic activity increases there is a concomitant deterioration in environmental quality. But beyond a turning point, as income increases the demand for a cleaner environment reduces the level of pollution.

Table 5 presents the results of Wald tests for the hypotheses that our three classes of variables are each jointly insignificant. Tests based on the entire US dataset are shown in column (1). [We discuss the other columns after presenting the regional estimates.] The data reject the null hypotheses that race/gender variables and economic variables do not influence toxic releases. The data fail to reject the null hypothesis that our set of political/ collective action variables does not influence releases, however. We next consider the individual coefficient estimates in Panel B of Table 4.

The impact of the variables with non-linear specifications depends on the level of the variables. Figure 1 illustrates the estimated impact for these

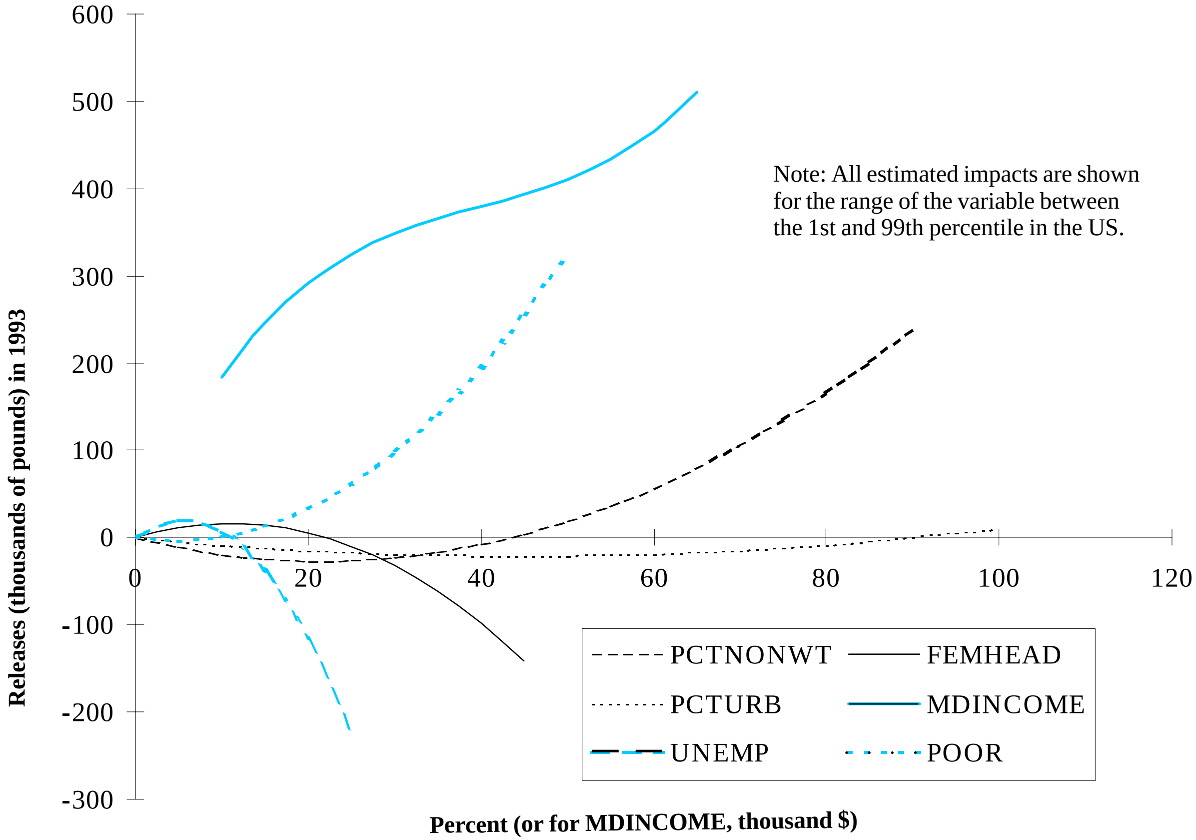
## Table 5. Wald Tests of Three Primary Hypotheses (Full Sample)

**Null Hypothesis: All variables in each group are jointly equal to zero**

| | Geographic Areas | | | |
|---|---|---|---|---|
| Variable Group | US<br>(1) | South<br>(2) | Non-South<br>(3) | California<br>(4) |
| Race/Gender<br>(5 variables) | 39.12<br>(<0.001) | 41.08<br>(<0.001)) | 6.94<br>(0.23) | 4.70<br>(0.70) |
| Economic<br>(9 variables) | 40.04<br>(<0.001) | 41.07<br>(<0.001) | 4.60<br>(0.87) | 5.22<br>(0.81) |
| Political/Collective Action<br>(8 variables) | 5.22<br>(0.73) | 11.13<br>(0.19) | 13.30<br>(0.10) | 11.33<br>(0.33) |

Notes: All test statistics are distributed as Chi-Squared under the null hypothesis (degrees of freedom equal to the number of variables indicated for each variable group, except for the California model). P-values indicated in parentheses. The number of restrictions for the California model (column 4) is 7 for Race/Gender, 9 for economic and 10 for political/collective action.

**Figure 1. Estimated Impact on 1993 Releases (Non.linear Variables)**



non-linear variables to aid in their interpretation.[18] In all cases the figure only displays the estimated impact for the range of the explanatory variable between the first and 99th percentile in the data. For example, we only display the impact of POOR below 50 percent, because the 99th percentile (across zipcodes) of the percentage of residents living in poverty is approximately 50 percent.

Consider first the race/gender variables. Releases are estimated to increase with the percentage of non-white population, once this percentage exceeds the turning point of approximately 22 percent. By contrast, releases generally fall with increases in the percentage of female-headed households, contrary to one possible view of environmental discrimination. Many of the economic variables also impact releases. Figure 1 shows that releases increase with increasing median household income. This general shape is not inconsistent with the inverse U-shaped environmental Kuznets' curve presented by Grossman and Krueger (1995) because of the variance in our parameter estimates.[19] Neighborhoods with a greater percentage of residents living in poverty (POOR) experience greater releases than less poverty-stricken neighborhoods. Finally, neighborhoods with high unemployment (above about 10 percent) experience fewer releases than low unemployment neighborhoods, as do neighborhoods with high residential vacancy rates (see Table 4 Panel B). These last two effects are due probably to generally depressed local economic conditions.

---

[18] Figures 1 and 2 are adjusted for the likelihood of a neighborhood experiencing any releases, from the Stage 1 models.

[19] For example, if the MEDINCSQ estimate fell by only -0.30 (only one-third of its standard error) to -2.71, the income-releases relationship would exhibit an inverse U-shape. We also explored the relationship between releases and median income, not controlling for all of the other demographic factors in our model. This is analogous to some reduced form estimates provided for developing countries. These estimates (not reported here) indicate a standard inverse U-shape, with a relatively low turning point at approximately the median income of $20,000.

## B. Southeastern US Estimates

The remaining columns of Table 4 present estimation results when segmenting the US into different regions. The estimates shown in column (3) are for the 11 southeastern states defined above, and the estimates shown in column (5) are for the remaining 39 states.[20] We were motivated to segment the US into geographic areas to capture potential regional differences influencing environmental outcomes.

Many parameter estimates differ in the two regions. In the South, the non-white population percentage significantly affects releases, while this variable is insignificant outside the South. Figure 2 illustrates that our model estimates for the South imply substantially higher releases for those neighborhoods with a large non-white population. [The Non-South estimated impact is shown for comparison, although this variable is not statistically significant in the Non-South dataset.] The other economic variables identified as significant at the 5-percent level in the full sample are also significant in the South sub-sample, with identical signs. These economic variables are insignificant in the Non-South sub-sample, however.

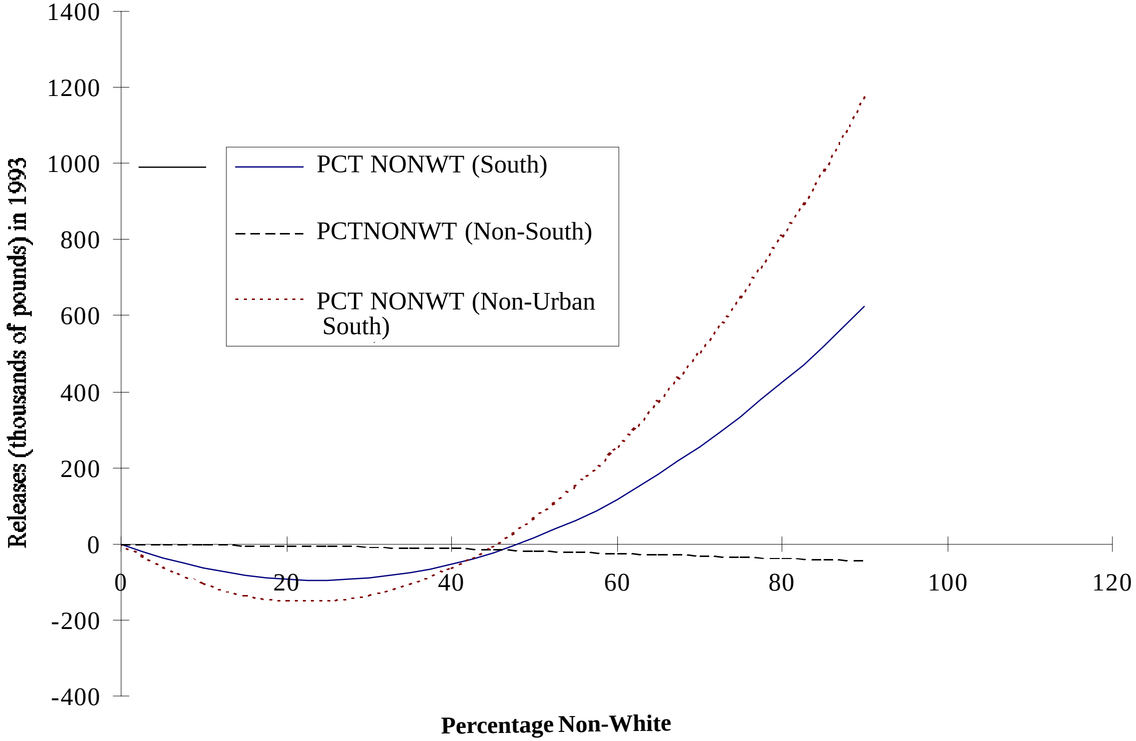
The Wald tests shown in columns (2) and (3) of Table 5 indicate that the southeastern US data reject the null hypotheses that the race/gender variables and the economic variables do not affect releases. The data do not reject the hypothesis that the set of political/collective action variables does not affect releases for the South. In the Non-South data, none of the three null hypotheses are rejected.

## C. California Estimates

The remaining results are based on the sub-sample of California zipcodes and are shown in column (7) of Table 4. This specification differs from the

[20] In the Non-South regression the omitted state is again California. In the South regression the omitted state is Florida. Nine of the ten remaining state dummies in the South regression are insignificant (Louisiana is significant and is 1148.8).

**Figure 2. Estimated Impact of Non-White Population Percentage: Southeastern U.S. vs. Non-South**



previous estimates in two ways. First, we specify the race variables slightly differently. As mentioned above, the correlation between the percentage of non-white residents and certain economic variables is substantial. For example, in the overall sample, the correlation coefficient between the percentage of non-white residents and the percentage of households living in poverty is 0.46. Fortunately, the data indicate that one minority group does not have this high correlation with economic characteristics: Asians. Unfortunately for our purposes, the percentage of Asian residents nationally is quite small, averaging 1.2 percent across zipcodes. This makes identifying an independent impact for this racial group unlikely based on the entire US sample.

However, the percentage of Asian residents is significantly greater in more racially diverse California, averaging 6.4 percent across zipcodes. This percentage also varies substantially across zipcodes in California and is uncorrelated with the percentage of residents living in poverty (the estimated correlation coefficient is -0.01). Therefore, the California specification in column (7) separates the non-white population percentage into two categories: percent Asian (PCTASIAN) and percent non-white and non-Asian (PCTNONWA). The results indicate whether an independent Asian effect is evident in the release data, and due to the nature of the data this effect is orthogonal to our poverty measures.

The second difference in the California estimates is the addition of two new variables: voter turnout (TURN90) and voting outcomes on Proposition 128 (PCT4_128), a wide-ranging initiative to improve environmental conditions. Voter turnout (defined as the percentage of residents that cast votes in the 1990 general election) ranged from 15 to 42 percent, with a mean of 28 and a median of 27 percent. The percentage of residents voting in favor of Proposition 128 ranged from 12 to 62 percent, with a mean of 33 and a median of 32 percent. As discussed above, these variables capture the political activity and environmental preferences of local residents.

Similar to the Non-South estimates, most of the variables in this model based only on California are insignificant. The key results from the California model are the following. First, the percentage of Asian residents as well as all other race/gender variables do not explain releases. Second, increased voter turnout has a negative but statistically insignificant impact on releases. Third, vote outcomes on proposition 128 have no impact on releases. The Wald tests based on California (column (4) of Table 5) indicate that none of the three joint null hypotheses are rejected.

**D. Non-Urban Estimates**

Due to land availability, population density and other factors, changes in release patterns may differ substantially between rural and urban areas.[21] The demographic composition of non-urban neighborhoods also varies considerably in different areas of the country. For example, as we document below, racial minorities represent a large portion of residents in some rural areas of the southeastern US, but elsewhere minority residents are more commonly concentrated in urban areas. If increases in toxic releases are more likely or less likely to be economically feasible in non-urban areas, the environmental impact on minority residents might differ across regions. The results previously presented in section IV. B. indicate that in the southeastern states, neighborhoods with a higher proportion of non-white residents are more likely to suffer from an increase in toxic releases. This section investigates whether this pattern could be due primarily to an increase in releases in non-urban areas, rather than to differences in neighborhood racial compositions. In particular, Table 6 reports estimates of the same models shown previously in Table 4, but for only non-urban zipcodes. We show that the key result concerning the concentration of minority residents is stronger when considering only non-urban zipcodes. This suggests that the increase in minority exposures in the South is not due mainly to an increase in releases in rural areas.

---

[21] We are grateful to the editor for encouraging us to investigate the changing release patterns of non-urban areas.

We exclude the predominantly urban zipcodes by dropping those in which more than 90 percent of the residents live in an "urban area."[22] The average population of the 23,354 zipcodes that satisfy this criterion is 4,671, compared to an average population of 23,306 for the 5,978 predominantly urban zipcodes. Non-white residents comprise more than 20 percent of the population in about 37 percent of the non-urban zipcodes in the South; by contrast, non-white residents comprise more than 20 percent of the population in only about 7 percent of the non-urban zipcodes outside the South. This discussion will focus on the Panel B results of Table 6, as well as the non-urban Wald test statistics reported in Table 7.

The results for the non-urban zipcodes are somewhat different from the full sample results. Consider first the race/gender variables. As in the full sample, the percentage of non-white residents affects releases primarily in the South. However, Figure 2 illustrates that the estimated increase in releases for predominantly non-white neighborhoods is more pronounced in southern, non-urban areas. In (unreported) estimates for *urban* zipcodes in the South, the percentage of non-white residents does not significantly affect releases. The evidence that minorities face increased exposures is therefore confined to non-urban areas of the South.

The second major difference in the non-urban sample is that many political/collective action variables are significantly different from zero. The Wald test statistics shown in Table 7 also indicate that this set of political/ collective action variables significantly affect releases in non-rural areas, contrary to the full sample tests shown in Table 5. In the South, surprisingly releases tend to be greater for non-urban neighborhoods that contain a greater fraction of households with children. The non-urban estimates for the South also indicate marginally significant impacts of the percentage of residents employed in manufacturing industries and the number of residents who

[22] For census purposes, an urbanized area consists of one or more places with a minimum population of 50,000 people plus adjacent area with a density of 1000 persons per square mile.

**Table 6. Estimation Results, by Region (Non-Urban Sample)**

**Panel A: Stage 1: Sample Selection Probit Equation (Dependent Variable is = 1 if Any Releases Reported)**

| | U.S. | | South | | Non-South | | California | |
|---|---|---|---|---|---|---|---|---|
| | (1) Estimate | (2) Std. Error | (3) Estimate | (4) Std. Error | (5) Estimate | (6) Std. Error | (7) Estimate | (8) Std. Error |
| Constant | -1.806*** | (0.262) | -0.891* | (0.540) | -2.235*** | (0.312) | 1.970 | (1.321) |
| FEMHEAD | 0.067*** | (0.004) | 0.079*** | (0.009) | 0.064*** | (0.006) | 0.017 | (0.029) |
| PCTFORN | 0.008* | (0.004) | 0.007 | (0.014) | 0.008* | (0.005) | 0.021 | (0.014) |
| PCTNONWT | -0.009*** | (0.001) | -0.008*** | (0.002) | -0.009*** | (0.002) | — | —- |
| PCTASIAN | — | — | — | — | — | — | -0.008 | (0.031) |
| PCTNONWA | — | — | — | — | — | — | -0.015 | (0.010) |
| VACANT | -0.002** | (0.001) | -0.003 | (0.003) | -0.003* | (0.001) | -0.020* | (0.011) |
| MDINCOME | 0.007** | (0.003) | 0.016** | (0.007) | 0.004 | (0.004) | -0.014 | (0.017) |
| POOR | -0.019*** | (0.003) | -0.021*** | (0.005) | -0.014*** | (0.003) | 0.008 | (0.022) |
| MEDROHU | -0.001*** | (0.000) | 0.000 | (0.000) | -0.001*** | (0.000) | 0.000 | (0.001) |
| UNEMP | -0.028*** | (0.005) | -0.009 | (0.010) | -0.039*** | (0.006) | 0.006 | (0.029) |
| BACH | -0.015*** | (0.002) | -0.025*** | (0.005) | -0.010*** | (0.003) | -0.003 | (0.013) |
| CARPOOL | -0.027*** | (0.003) | -0.035*** | (0.005) | -0.022*** | (0.003) | 0.002 | (0.018) |
| HHWKIDS | -0.002 | (0.003) | -0.013** | (0.006) | -0.001 | (0.003) | -0.025** | (0.012) |
| MANU | 0.036*** | (0.001) | 0.028*** | (0.002) | 0.044*** | (0.002) | 0.044*** | (0.013) |
| MEDAGE | -0.002 | (0.004) | -0.014* | (0.009) | 0.002 | (0.005) | -0.059** | (0.027) |
| RENTPCT | 0.013*** | (0.002) | 0.009** | (0.004) | 0.015*** | (0.002) | 0.000 | (0.009) |
| TURN90 | — | — | — | — | — | — | -0.011 | (0.014) |
| PCT4_128 | — | — | — | — | — | — | -0.022 | (0.015) |
| TOTPOP | 0.070*** | (0.002) | 0.077*** | (0.004) | 0.067*** | (0.003) | 0.051*** | (0.008) |
| PCTURB | 0.009*** | (0.001) | 0.005*** | (0.001) | 0.010*** | (0.001) | 0.008** | (0.003) |
| Log-Likelihood | | -8018 | | -2264 | | -5710 | | -194 |

**Table 6 (continued). Estimation Results, by Region (Non-Urban Sample)**

**Panel B: Stage 2: Dependent Variable is 1993 Releases (in Thousands of Pounds)**

| | U.S. | | South | | Non-South | | California | |
|---|---|---|---|---|---|---|---|---|
| | (1) Estimate | (2) Std. Error | (3) Estimate | (4) Std. Error | (5) Estimate | (6) Std. Error | (7) Estimate | (8) Std. Error |
| Constant | -5915.00*** | (1051.70) | -9530.20*** | (2179.10) | -1164.30 | (1021.10) | 2367.20 | (2002.20) |
| *Race/Gender Variables* | | | | | | | | |
| FEMHEAD | -19.54 | (30.12) | -56.82 | (79.76) | -5.55 | (25.70) | -211.67* | (124.59) |
| FEMHEDSQ | -0.42 | (0.78) | -0.56 | (1.80) | -0.24 | (0.84) | 7.25* | (3.89) |
| PCTFORN | 7.73 | (13.55) | 31.99 | (48.02) | -15.36 | (10.04) | -11.87 | (21.98) |
| PCTNONWT | -41.17*** | (7.81) | -63.16*** | (14.88) | -1.25 | (7.73) | — | —- |
| PCTNWTSQ | 0.98*** | (0.11) | 1.39*** | (0.21) | 0.09 | (0.12) | — | —- |
| PCTASIAN | — | — | — | — | — | — | -35.22 | (73.38) |
| PCTASQ | — | — | — | — | — | — | 3.37 | (4.59) |
| PCTNONWA | — | — | — | — | — | — | -33.68 | (22.68) |
| PCTNWASQ | — | — | — | — | — | — | 0.18 | (0.25) |
| *Economic Variables* | | | | | | | | |
| VACANT | -5.59 | (3.72) | -25.95** | (13.22) | -2.46 | (2.43) | -40.58** | (16.85) |
| MDINCOME | 237.27*** | (60.39) | 424.10* | (227.91) | -14.02 | (45.79) | 126.08 | (236.75) |
| MEDINCSQ | -5.98*** | (1.68) | -10.64 | (8.10) | 0.34 | (1.23) | -4.91 | (6.78) |
| MEDINCCU | 0.05*** | (0.01) | 0.08 | (0.09) | 0.00 | (0.01) | 0.06 | (0.06) |
| POOR | -14.42 | (19.26) | -33.95 | (53.78) | -2.96 | (13.99) | 189.39*** | (71.96) |
| POORSQ | 1.10** | (0.45) | 2.36** | (1.05) | 0.21 | (0.36) | -5.68*** | (2.06) |
| MEDROHU | -0.81 | (0.57) | -0.71 | (1.53) | -0.11 | (0.43) | 1.25 | (1.28) |
| UNEMP | 119.30*** | (36.68) | 270.45*** | (93.79) | 14.72 | (29.17) | -122.03 | (77.13) |
| UNEMPSQ | -9.45*** | (2.01) | -21.69*** | (4.90) | -0.97 | (1.66) | 5.13* | (2.65) |

**Table 6 (continued). Panel B**

| | U.S. | | South | | Non-South | | California | |
|---|---|---|---|---|---|---|---|---|
| | (1) Estimate | (2) Std. Error | (3) Estimate | (4) Std. Error | (5) Estimate | (6) Std. Error | (7) Estimate | (8) Std. Error |
| *Political/Collective Action Variables* | | | | | | | | |
| BACH | -1.73 | (6.75) | 6.22 | (18.26) | -9.57** | (4.75) | -28.86* | (16.55) |
| CARPOOL | 68.11** | (28.17) | 151.57* | (85.98) | 4.54 | (19.48) | -129.23*** | (43.09) |
| CARPOLSQ | -1.59** | (0.73) | -3.48* | (2.00) | 0.31 | (0.54) | 2.73*** | (0.88) |
| HHWKIDS | 104.13*** | (34.54) | 184.95** | (71.94) | 40.18 | (32.91) | 49.91 | (127.74) |
| HHWKIDSQ | -1.15*** | (0.36) | -2.23*** | (0.79) | -0.45 | (0.33) | -0.24 | (1.23) |
| MANU | -10.97 | (6.86) | -24.13* | (13.15) | -2.28 | (5.95) | -22.08 | (27.04) |
| MEDAGE | 17.38 | (13.10) | 7.19 | (32.99) | 11.69 | (10.60) | -66.98 | (53.32) |
| RENTPCT | 2.13 | (5.75) | 18.64 | (14.91) | 4.78 | (4.22) | -8.75 | (12.22) |
| TURN90 | — | — | — | — | — | — | -19.77 | (14.73) |
| PCT4_128 | — | — | — | — | — | — | 5.69 | (21.91) |
| *ControlVariables* | | | | | | | | |
| TOTPOP | 11.05 | (7.61) | 7.27 | (18.81) | 11.01** | (5.37) | 12.68 | (21.98) |
| 1990 Releases | 1.00*** | (0.01) | 1.17*** | (0.01) | 0.62*** | (0.01) | 0.21*** | (0.06) |
| PCTURB | -2.04 | (4.44) | -4.85 | (10.26) | 1.98 | (3.22) | -3.15 | (13.34) |
| PCTURBSQ | 0.03 | (0.05) | -0.05 | (0.12) | -0.01 | (0.03) | 0.04 | (0.11) |
| Inverse Mill's Ratio | 120.87 | (228.56) | -128.31 | (507.54) | 95.25 | (167.97) | 725.05 | (768.21) |
| Number of Observations | | 23354 | | 5609 | | 17745 | | 723 |
| Adjusted $R^2$ | | 0.755 | | 0.835 | | 0.579 | | 0.535 |

Notes: * denotes significantly different from zero at the 10 percent level; ** denotes significantly different from zero at the 5 percent level; *** denotes significantly different from zero at the 1 percent level (all two-tailed tests).

**Table 7. Wald Tests of Three Primary Hypotheses (Non-Urban Sample)Z**

**Null Hypothesis: All variables in each group are jointly equal to zero**

| | Geographic Areas | | | |
|---|---|---|---|---|
| Variable Group | US (1) | South (2) | Non-South (3) | California (4) |
| Race/Gender | 100.70 | 62.38 | 3.50 | 12.92 |
| (5 variables) | (<0.001) | (<0.001)) | (0.62) | (0.07) |
| Economic | 44.76 | 45.40 | 3.45 | 19.64 |
| (9 variables) | (<0.001) | (<0.001) | (0.94) | (0.02) |
| Political/Collective Action | 27.07 | 18.89 | 19.04 | 39.46 |
| (8 variables) | (<0.001) | (0.02) | (0.01) | (<0.001) |

Notes: All test statistics are distributed as Chi-Squared under the null hypothesis (degrees of freedom equal to the number of variables indicated for each variable group, except for the California model). P-values indicated in parentheses. The number of restrictions for the California model (column 4) is 7 for Race/Gender, 9 for economic and 10 for political/collective action.

carpool. The non-urban estimates for the non-southern states (column 5 of Table 6, Panel B) indicate that releases are lower in neighborhoods with a higher percentage of adults with bachelor's degrees. Finally, the non-urban estimates for California indicate that releases are lower in neighborhoods where a higher percentage of workers use carpools.

In summary, these estimates based on only non-urban zipcodes suggest that residents in predominantly non-white, southern rural areas were exposed to more toxic releases than their urban counterparts. The results fail to support the hypothesis that the greater releases estimated for southern non-white neighborhoods are due mainly to the greater releases that may occur in non-urban areas. The results also indicate that our political/collective action variables have a greater influence on releases in non-urban areas, which is an intriguing finding that warrants future study.

**E. Alternative Specifications**

In this subsection we briefly discuss several alternative model specifications, although we do not report them in detail here in order to conserve space.

The TRI reports transfers (or "shipments") of toxic chemicals, which are typically directed toward publicly owned treatment works (POTW). The accounting of these transfers has been more accurate than the accounting of releases, at least in the early years of the TRI. In recent years these off-site transfers have been growing dramatically. For example, while toxic releases fell by 6.5 percent between 1990 and 1993 (see Table 3), toxic transfers increased by more than 200 percent—from 1.16 billion to 3.86 billion pounds. While this reflects an overall increase in the generation of toxic chemicals, these transfers remove the toxic chemicals from the local environment and are often associated with reduced local environmental releases. Consequently, increases in transfers often improve the local environment, unlike increases in releases.

We were unable to find strong evidence that transfers are closely related to the demographic and economic characteristics of the zipcode neighborhood surrounding manufacturing facilities. We estimated a set of sample selection models similar to those shown in Table 4, except with 1993 transfers replacing releases as the dependent variable (and 1990 transfers replacing releases as a control explanatory variable). The overall fit of the models was poor, as reflected in adjusted R-square statistics that were below 0.01 for the entire U.S., the South and the Non-South datasets. Individual coefficient estimates were significantly different from zero only rarely.[23]

We also investigated whether systematic initial underreporting or overreporting of releases might be able to explain our finding that releases tended to increase between 1990 and 1993 in non-urban, southern zipcodes with a high proportion of non-white residents. Some small firms initially may have failed to comply with reporting requirements.[24] If these underreported releases varied systematically by region (and with demographic or economic characteristics of the zipcodes), then our results could be biased.

To reduce any bias due to underreporting, we divided facilities into three classes: (1) those with positive releases or transfers reported in both 1990 and 1993; (2) those with positive releases or transfers only in 1990 (but no data reported in 1993); and, (3) those with positive releases or transfers only in 1993 (but no data reported in 1990). This last group might be non-reporting

[23] We also estimated a model with total 1993 releases and transfers as the dependent variable—which is a measure of overall toxic chemical "generation" in the zipcode. The demographic and economic characteristics in this model can explain some of the variation in generation across zipcodes (e.g., the adjusted R-square is 0.33 for the entire U.S. dataset); however, the coefficient estimates are difficult to interpret because—as discussed above—increases in releases can harm the local environment while increases in transfers can improve the local environment.

[24] Brehm and Hamilton (1996) find that in Minnesota, small firms that generated small amounts of toxic chemicals were most likely to fail to file TRI reports in 1991. They attribute such noncompliance to ignorance rather than (strategic) evasion of the law. We are grateful to an anonymous referee for suggesting that we study the impact of under and over reporting.

in 1990, and by 1993 they had begun to comply with the TRI reporting requirements.[25]

We estimated the same models reported above on only the facilities in group (1) (i.e., those reporting in both years), to determine if our main conclusions continue to hold on a dataset with less potential bias from underreporting. Our conclusions tend to be somewhat weaker, but they hold up qualitatively. For the full U.S. dataset, the percentage of non-white residents does not significantly affect releases, although this variable continues to affect releases significantly in the southeastern states estimates. The main difference in the results for this sub-sample of facilities is that the percentage of residents who use carpools significantly affects releases, and this makes the political/collective action Wald test statistics significant in the entire U.S. estimates as well as the estimates for the southeastern states.

Finally, we also reestimated the models after disaggregating releases by pollution media. Our main results in Table 4 are based on total releases, which include releases to air, surface water, underground injections and land. It is possible that race, economic and collective action influences affect these kinds of releases differently, due perhaps to public and regulator scrutiny that differs depending on the type of pollution. About 45 percent of releases are to air, so not surprisingly the air release estimates generally parallel those in Table 4. The main difference is that median income is not significant in any of the air release estimates. In addition, in the air release model estimated for the southeastern states, the estimated impact of the percentage of non-white residents is much smaller in magnitude, although it remains statistically significant. Water and land releases represent about 18 percent and about 8 percent of the total releases, but our set of economic and demographic characteristics fail to explain releases in these media. Individual coefficients are rarely statistically different from zero in any estimates. Finally, facilities

[25] Of course, many of the facilities in group (3) are new facilities that began releasing toxic chemicals between 1990 and 1993, and many of the facilities in group (2) were closed between 1990 and 1993.

release toxic chemicals by underground injection in only about 1 percent of the zipcodes, although by weight, releases of this type represent about 29 percent of the total. The small number of zipcodes experiencing underground releases leads to unreliable or unsuccessful estimation results for the sample selection model.

## V. Summary

This paper presents a reduced form statistical analysis of the relationship between environmental outcomes and neighborhood characteristics throughout the United States. We also conduct regional regressions within the United States to capture differences across geographic areas. Our approach uses the level of toxic chemical releases in 1993 as the measure of environmental performance, based on the Toxics Release Inventory, and we control for 1990 releases. The 1990 US Census provides the data on neighborhood characteristics, and the analysis is conducted at the zipcode level. The goal is to distinguish between three alternative explanations for differences in environmental outcomes—race/gender influences, an economic (Coasian) explanation, and an explanation based on political/collective action.

Many economic variables significantly impact releases for the overall sample and within the southeastern states. The estimates based on the entire US indicate that releases increase as income increases, but the imprecision of our estimates does not rule out the possible existence of an inverse U-shaped Environmental Kuznets' curve (i.e., a reduction in releases with increasing income once income exceeds some threshold). Releases also tend to by lower in areas with high unemployment rates.

While the scope of our inquiry was much broader than a simple search for environmental injustice, our most provocative finding is that race appears to be an important determinant of releases in the South. This result seems confined to non-urban areas, which contain high concentrations of minority residents mainly in the South. This pattern of increased releases in minority areas controls for many other economic and collective action variables, and it is not observed outside the South or in predominantly urban areas. This finding has important implications for the debate on environmental equity, and is consistent with case study evidence.[26]

While our study provides some statistical evidence of environmental inequity, it also provides some evidence of the potential power of collective action. Outside of the urban centers and also most prominently in the southeast, the variables that proxy political/collective action incentives and preferences significantly influence environmental outcomes. This suggests that raising awareness and providing information to the affected rural, southern communities may be a significant step in improving environmental performance in areas that may suffer from environmental injustice.

## References

Anderton, D. L., A. B. Anderson, J. M. Oakes, M. R. Fraser (1994), "Environmental Equity: The Demographics of Dumping," *Demography* 31, pp. 229-248.

Arora, Seema and Timothy N. Cason (1995), "An Experiment in Voluntary Environmental Regulation: Participation in EPA's 33/50 Program," *Journal of Environmental Economics and Management* 28, pp. 271-286.

Been, Vicki (1994), "Unpopular Neighbors: Are Dumps and Landfills Sited Equitably?" *Resources*, no. 115 (Spring), pp. 16-19.

Brehm, John and James T. Hamilton (1996), "Noncompliance in Environmental Reporting: Are Violators Ignorant, or Evasive, of the Law?" *American Journal of Political Science* 40, pp. 444-477.

Breusch, T. and A. Pagan (1980), "The LM Test and its Application to Model Specification in Econometrics," *Review of Economic Studies* 47, pp. 239-254.

[26] Our findings echo the tales of Afton and Warren counties in North Carolina documented in Bullard (1990).

Bryant, B. and P. Mohai, eds. (1992), *Race and the Incidence of Environmental Hazards: A Time for Discourse* (Boulder CO: Westview Press).

Bullard, Robert D. (1983), "Solid Waste Sites and the Black Houston Community," *Sociological Inquiry* 53, pp. 273-288.

Bullard, Robert D. (1990), *Dumping In Dixie: Race, Class, and Environmental Quality* (Boulder CO: Westview Press).

Filer, John E., Lawrence W. Kenny and Rebecca B. Morton (1993), "Redistribution, Income and Voting," *American Journal of Political Science* 37, pp. 63-87.

Glickman, Theodore S. and Robert Hersh (1995) "Evaluating Environmental Equity: The Impact of Health Hazards on Selected Groups in Allegheny County, Pennsylvania," Resources for the Future Discussion paper No. 95-13.

Goldman, Benjamin A. and L. Fritton (1994), "Toxic Wastes and Race Revisited," Center for Policy Alternatives, National Association for the Advancement of Colored People, United Church of Christ Commission for Racial Justice.

Grossman, Gene M. and Alan B. Krueger (1995), "Economic Growth and the Environment," *Quarterly Journal of Economics*, 110(2), pp. 353-377.

Hamilton, James H. (1995a), "Testing for Environmental Racism: Prejudice, Profits, Political Power?", *Journal of Policy Analysis and Management* 14, pp. 107-132.

Hamilton, James H. (1995b), "Pollution as News: Media and Stock Reactions to the Toxics Release Inventory Data," *Journal of Environmental Economics and Management*, 28, pp. 98-113.

Heckman, James (1979), "Sample Selection Bias as a Specification Error," *Econometrica* 47, pp. 153-162.

Kahn, Matthew, E. and John G. Matsusaka (1997), "Demand for Environmental Goods: Evidence from Voting Patterns on California Initiatives," *Journal of Law and Economics* 40, pp. 137-173.

Konar, Shameek and Mark A. Cohen (1997), "Information as Regulation: The Effect of Community Right to Know Laws on Toxic Emissions," *Journal of Environmental Economics and Management* 32, pp. 109-124.

Olson, Mancur (1965), *The Logic of Collective Action* (Cambridge, MA: Harvard University Press).

Scholz, John, T. and Wayne B. Gray (1997), "Can Government Facilitate Cooperation? An Informational Model of OSHA Enforcement," *American Journal of Political Science* 41, pp. 693-717.

U.S. Environmental Protection Agency (1989), Toxic Chemical Release Inventory Risk Screening Guide, Volume 2: Appendices, Office of Toxic Substances, EPA 560/2-89-002, July.

U.S. Environmental Protection Agency (1992), Toxic Chemical Release Inventory: Reporting Form R and Instructions, Revised 1991 Version, 700-K-92-002, May.

U.S. General Accounting Office (1983), "Siting of Hazardous Waste Landfills and Their Correlation with Racial and Economic Status of Surrounding Communities," GAO 121648.

**Appendix. Ordinary Least Squares Estimation Results, by Region (Full Sample)**

**Dependent Variable is 1993 Releases (in Thousands of Pounds)**

| | U.S. | | South | | Non-South | | California | |
|---|---|---|---|---|---|---|---|---|
| | (1) Estimate | (2) Std. Error | (3) Estimate | (4) Std. Error | (5) Estimate | (6) Std. Error | (7) Estimate | (8) Std. Error |
| Constant | -1753.30*** | (486.61) | -5897.10*** | (1502.10) | 314.36 | (365.89) | 671.96 | (1537.80) |
| | *Race/Gender Variables* | | | | | | | |
| FEMHEAD | 8.93 | (12.65) | 11.41 | (39.75) | -4.51 | (9.55) | -61.28 | (58.87) |
| FEMHEDSQ | -0.63*** | (0.21) | -1.19** | (0.60) | -0.02 | (0.16) | 1.08 | (1.13) |
| PCTFORN | -6.90 | (5.44) | 3.32 | (18.97) | -6.41 | (3.96) | -12.36 | (12.51) |
| PCTNONWT | -10.40** | (4.76) | -39.66*** | (11.58) | 0.50 | (3.95) | | |
| PCTNWTSQ | 0.27*** | (0.06) | 0.79*** | (0.14) | -0.02 | (0.05) | | |
| PCTASIAN | — | — | | | | | -10.71 | (28.19) |
| PCTASQ | — | — | | | | | 0.21 | (0.66) |
| PCTNONWA | — | — | | | | | -6.91 | (18.17) |
| PCTNWASQ | — | — | | | | | -0.05 | (0.20) |
| | *Economic Variables* | | | | | | | |
| VACANT | -5.07* | (2.93) | -22.77* | (11.95) | -2.16 | (1.94) | -13.30 | (15.75) |
| MDINCOME | 110.56*** | (34.46) | 362.31*** | (138.39) | -11.71 | (24.57) | 109.48 | (141.25) |
| MEDINCSQ | -2.46*** | (0.91) | -8.33* | (4.37) | 0.28 | (0.63) | -3.16 | (3.94) |
| MEDINCCU | 0.02** | (0.01) | 0.06 | (0.04) | 0.00 | (0.01) | 0.03 | (0.03) |
| POOR | -6.89 | (11.87) | -48.99 | (42.77) | 5.32 | (8.22) | 45.07 | (47.89) |
| POORSQ | 0.77** | (0.24) | 2.37*** | (0.77) | -0.08 | (0.18) | -0.90 | (0.95) |
| MEDROHU | -0.54 | (0.36) | -1.06 | (1.33) | -0.17 | (0.25) | 0.07 | (1.05) |
| UNEMP | 37.60* | (21.01) | 232.09*** | (75.37) | 5.49 | (15.37) | -96.23 | (99.10) |
| UNEMPSQ | -3.04*** | (0.92) | -15.96*** | (3.72) | -0.22 | (0.67) | 3.27 | (3.56) |

## Appendix (continued)

| | U.S. | | South | | Non-South | | California | |
|---|---|---|---|---|---|---|---|---|
| | (1) Estimate | (2) Std. Error | (3) Estimate | (4) Std. Error | (5) Estimate | (6) Std. Error | (7) Estimate | (8) Std. Error |
| *Political/Collective Action Variables* | | | | | | | | |
| BACH | 0.39 | (3.75) | 6.12 | (13.10) | -4.82* | (2.59) | -19.66 | (12.39) |
| CARPOOL | 26.38 | (20.28) | 93.49 | (70.55) | 0.69 | (14.16) | -36.43 | (53.35) |
| CARPOLSQ | -0.48 | (0.54) | -2.18 | (1.70) | 0.29 | (0.40) | 0.62 | (0.98) |
| HHWKIDS | 5.96 | (20.30) | 37.42 | (57.19) | -3.89 | (15.29) | -14.69 | (53.78) |
| HHWKIDSQ | -0.10 | (0.21) | -0.54 | (0.61) | 0.01 | (0.16) | 0.35 | (0.52) |
| MANU | -8.19*** | (2.82) | -16.44** | (7.85) | -1.70 | (2.17) | 5.44 | (11.62) |
| MEDAGE | 9.52 | (7.95) | 3.13 | (25.99) | 3.97 | (5.71) | 5.69 | (27.34) |
| RENTPCT | 3.40 | (2.92) | 18.38* | (10.21) | 1.85 | (2.03) | 10.11 | (9.25) |
| TURN90 | — | — | | | | | -17.91 | (15.13) |
| PCT4_128 | — | — | | | | | 11.72 | (12.24) |
| *Control Variables* | | | | | | | | |
| TOTPOP | 1.58 | (2.28) | 0.67 | (7.66) | 3.83** | (1.56) | 0.76 | (4.42) |
| 1990 Releases | 0.94*** | (0.01) | 1.07*** | (0.01) | 0.61*** | (0.01) | 0.34*** | (0.08) |
| PCTURB | -6.68*** | (2.27) | -9.29 | (6.82) | -0.10 | (1.62) | -26.54** | (11.91) |
| PCTURBSQ | 0.06** | (0.02) | 0.01 | (0.07) | 0.01 | (0.02) | 0.22* | (0.12) |
| Number of Observations | | 8121 | | 2052 | | 6069 | | 474 |
| Adjusted $R^2$ | | 0.735 | | 0.813 | | 0.529 | | 0.042 |

Notes: * denotes significantly different from zero at the 10 percent level; ** denotes significantly different from zero at the 5 percent level; *** denotes significantly different from zero at the 1 percent level (all two-tailed tests).